



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/611,368	07/01/2003	Hua Wu	M61.12-0510	3170
27366	7590	06/12/2007	EXAMINER	
WESTMAN CHAMPLIN (MICROSOFT CORPORATION) SUITE 1400 900 SECOND AVENUE SOUTH MINNEAPOLIS, MN 55402-3319			NEWAY, SAMUEL G	
ART UNIT		PAPER NUMBER		
2626				
MAIL DATE		DELIVERY MODE		
06/12/2007		PAPER		

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary	Application No.	Applicant(s)
	10/611,368	WU ET AL.
	Examiner Samuel G. Neway	Art Unit 2626

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1) Responsive to communication(s) filed on 01 July 2003.
 2a) This action is **FINAL**. 2b) This action is non-final.
 3) Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4) Claim(s) 1-22 is/are pending in the application.
 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
 5) Claim(s) _____ is/are allowed.
 6) Claim(s) 1-3,6-15 and 18-22 is/are rejected.
 7) Claim(s) 4,5,16 and 17 is/are objected to.
 8) Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9) The specification is objected to by the Examiner.
 10) The drawing(s) filed on 01 July 2003 is/are: a) accepted or b) objected to by the Examiner.
 Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
 Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
 11) The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12) Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
 a) All b) Some * c) None of:
 1. Certified copies of the priority documents have been received.
 2. Certified copies of the priority documents have been received in Application No. _____.
 3. Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) <input checked="" type="checkbox"/> Notice of References Cited (PTO-892)	4) <input type="checkbox"/> Interview Summary (PTO-413)
2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948)	Paper No(s)/Mail Date. _____
3) <input type="checkbox"/> Information Disclosure Statement(s) (PTO/SB/08)	5) <input type="checkbox"/> Notice of Informal Patent Application
Paper No(s)/Mail Date _____. 	6) <input type="checkbox"/> Other: _____

DETAILED ACTION

1. This is responsive to the Application filed on 1 July 2003.

Drawings

2. The drawings are objected to because in FIG. 2B, item 216, the word "INDOCATOR" is believed to be a typographical error. Corrected drawing sheets in compliance with 37 CFR 1.121(d) are required in reply to the Office action to avoid abandonment of the application. Any amended replacement drawing sheet should include all of the figures appearing on the immediate prior version of the sheet, even if only one figure is being amended. The figure or figure number of an amended drawing should not be labeled as "amended." If a drawing figure is to be canceled, the appropriate figure must be removed from the replacement sheet, and where necessary, the remaining figures must be renumbered and appropriate changes made to the brief description of the several views of the drawings for consistency. Additional replacement sheets may be necessary to show the renumbering of the remaining figures. Each drawing sheet submitted after the filing date of an application must be labeled in the top margin as either "Replacement Sheet" or "New Sheet" pursuant to 37 CFR 1.121(d). If the changes are not accepted by the examiner, the applicant will be notified and informed of any required corrective action in the next Office action. The objection to the drawings will not be held in abeyance.

Claim Rejections - 35 USC § 102

3. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

4. Claims 1, 10 – 15, and 18 – 22 are rejected under 35 U.S.C. 102(b) as being anticipated by Gasperin et al (“Using Syntactic Contexts for Measuring Word Similarity”, Workshop on Knowledge Acquisition and Categorization, ESSLLI (2001)).

Claim 1:

Gasperin discloses a method of determining similarity between words, comprising:

receiving as an input a first word and a first dependency structure (attribute) that includes the first word; receiving a data structure indicative of a second word and a second dependency structure that includes the second word (“calculates the similarity value between two words, m and n, by comparing the attributes they share”, section 3.1. It is inherent that the words are received before they are compared); and

calculating the similarity between the first and second words based on a similarity measure (“calculates the similarity value between two words”, section 3.1) weighted based on a frequency indicator indicative of a frequency of occurrence of the second dependency structure in training data (“The local weight l_w is based on the frequency of the attribute with a given word”, section 3.1).

Claim 10:

Gasperin discloses a method of generating annotated data for use in calculating similarity between words, comprising:

receiving a textual input; parsing the textual input into dependency structures including words and relation types indicative of relations between the words in the textual input (“... each word in the corpus is associated to a set of syntactic contexts.”, section 1, paragraph 2);

generating a vector corresponding to each dependency structure, the vector including a related word, a relation type indicator, and a frequency indicator indicating a frequency with which the dependency structure occurred in the textual input; and storing the words and corresponding vectors regardless of the frequency with which the dependency structures occurred in the textual input (section 1, paragraph 2).

Claim 11:

Gasperin discloses the method of claim 10 wherein the frequency indicator comprises a normalized count value (“The local weight lw is based on the frequency of the attribute with a given word”, section 3.1. Note that lw is calculated using a log function that normalizes the count value).

Claim 12:

Gasperin discloses the method of claim 10 wherein parsing the textual input into dependency structures comprises: parsing the textual input into dependency triples (“ $<R, w1, w2>$ ”, section 2.1).

Claim 13:

Gasperin discloses a natural language processing system, comprising:

a data store storing head words and associated attributes, each of the attributes including a related word that was related to the head word in a training corpus, a relation type indicator indicating a type of relation between the head word and the related word ("<R, w1, w2>", section 2.1), and a frequency indicator indicative of a frequency with which the attribute occurred relative to the head word in the training corpus ("The local weight l_w is based on the frequency of the attribute with a given word", section 3.1); and

a similarity generator configured to receive an input word and an associated input dependency structure and to access the data store and calculate a similarity between the input word and head words in the data structure based on the input word and associated input dependency structure and the head words and associated dependency structures using a similarity measure that weights a similarity corresponding to a given head word based on the frequency indicator associated with the given word ("calculates the similarity value between two words", section 3.1).

Claim 14:

Gasperin discloses the system of claim 13 wherein the similarity generator is configured to weight the similarity with a weighting measure ("3.1 The Weighted Jaccard Similarity Measure", Title of section 3.1).

Claim 15:

Gasperin discloses the system of claim 14 wherein the similarity generator is configured to select one of a plurality of weighting measures to weight the similarity based on the frequency indicator associated with the given head word.

Claim 18:

Gasperin discloses the system of claim 13 and further comprising: a lexical knowledge base, the similarity generator being configured to access the lexical knowledge base to identify a subset of the head words in the data store as candidate words prior to calculating the similarity (“Roget’s thesaurus, WordNet, … ”, section 1, paragraph 1).

Claim 19:

Gasperin discloses the system of claim 13 wherein the data store stores the attributes as vectors (“… each word in the corpus is associated to a set of syntactic contexts”, section 1, paragraph 2).

Claims 20 – 22:

Claims 20 – 22 are similar in scope and content to claims 10 – 12 and are rejected with the same rationale.

Claim Rejections - 35 USC § 103

5. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

6. Claims 2 – 3, and 6 – 9 are rejected under 35 U.S.C. 103(a) as being unpatentable over Gasperin et al ("Using Syntactic Contexts for Measuring Word Similarity", Workshop on Knowledge Acquisition and Categorization, ESSLLI, 2001) in view of Evert et al ("Methods for the Qualitative Evaluation of Lexical Association Measures", Proceedings of the 39th Annual Meeting of Association for Computational Linguistics, 2001).

Claims 2 – 3:

Gasperin discloses the method of claim 1 but it does explicitly disclose using a plurality of weighting measures including a co-occurrence weighting measure and a mutual information (MI) weighting measure.

Evert discloses lexical measures similar to Gasperin's wherein these measures include co-occurrence weighting measures and mutual information (MI) weighting measures ("The measures – Mutual Information (MI), ... and co-occurrence", section 1, paragraph 3).

It would have been obvious to one with ordinary skill in the art at the time of the invention to use a plurality of weighting measures in Gasperin's method because "In computational linguistics, a variety of ... measures have been proposed" (Evert, section 1, paragraph 1).

Claim 6:

Gasperin and Evert disclose the method of claim 2, Gasperin further discloses wherein receiving a data structure indicative of a second word comprises: accessing a

data store that stores records that include words and associated dependency structures and frequency indicators (section 1, paragraph 2).

Claim 7:

Gasperin and Evert disclose the method of claim 6, Gasperin further discloses wherein the associated dependency structures and frequency indicators in the data store are stored as vectors associated with the words, and wherein accessing a data store comprises: accessing the words and associated vectors (section 1, paragraph 2).

Claim 8:

Gasperin and Evert disclose the method of claim 6, Gasperin further discloses wherein accessing the data store comprises: identifying candidate words in the data store by reducing the search space of records in the data store (“Roget’s thesaurus, WordNet, ... ”, section 1, paragraph 1).

Claim 9:

Gasperin and Evert disclose the method of claim 8 Gasperin further discloses wherein identifying candidate words comprises: accessing a lexical knowledge base to identify possible candidate words in the data store (“Roget’s thesaurus, WordNet, ... ”, section 1, paragraph 1).

Allowable Subject Matter

7. Claims 4 – 5 and 16 – 17 are objected to as being dependent upon a rejected base claim, but would be allowable if rewritten in independent form including all of the limitations of the base claim and any intervening claims.

These claims are allowable because the prior art of record does not teach using either the co-occurrence frequency measure or the MI measure depending on the frequency of occurrence as claimed in the instant claims.

Conclusion

8. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

Richardson et al (USPN 6,098,033) discloses a facility for determining similarity between two input words utilizing the frequencies with which path patterns occurring between the words occur between words known to be synonyms.

Messerly et al (USPN 6,076,051) discloses a method performing information retrieval utilizing semantic representation of text.

Dolan et al (USPN 6,871,174) discloses a natural language processing system to determine a relationship (such as similarity in meaning) between two textual segments.

Corston-Olivier et al (USPN 6,901,402) discloses a system for determining a relationship (such as similarity in meaning) between two or more textual inputs.

9. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Samuel G. Neway whose telephone number is 571-270-1058. The examiner can normally be reached on Monday - Friday 8:30AM - 5:30PM EST.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, David R Hudspeth can be reached on 571-272-7843. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.


DAVID HUDSPETH
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 2600

SN

